CHAPTER 19 Sampling

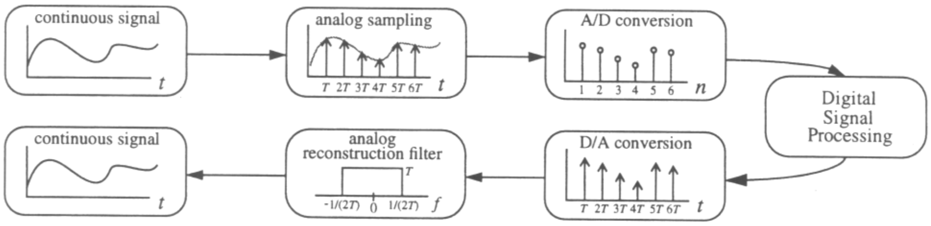
Overview

Sampling theory has become more and more critical in the "digital age." Most signals that we work with are continuous, but they are increasingly being processed using digital computers. Sampling theory gives us the tools to faithfully convert from the analog world to digital and back again. The fundamental result of this chapter is that any bandlimited signal can be *completely* characterized by discrete, equally spaced samples, provided that the samples are spaced close enough together in time.

19.1 What is Sampling?

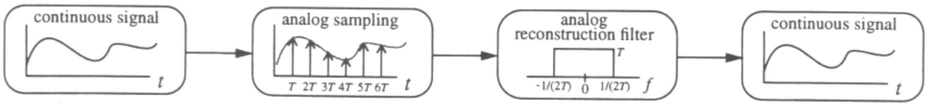
Sampling is the process of taking a continuous-time signal and representing it by a series of discrete samples. Reconstruction is the process of taking these discrete samples and recreating the associated continuous-time signal. These processes are illustrated in the following block diagram.

Overall process:



We will leave the processes of A/D, D/A, and digital signal processing for another book. Why do we need sampling theory at all? We could just work with analog signals only. Doing signal processing in the digital world, i.e. on computers, provides far more flexibility, but with the expense of added complexity. Filter characteristics are easily changed by reprogramming a few numbers instead of having to change resistors and capacitors. Also, the processing is essentially noise-free, unlike the "fuzz" that you sometimes see on your oscilloscopes when doing analog circuit design.

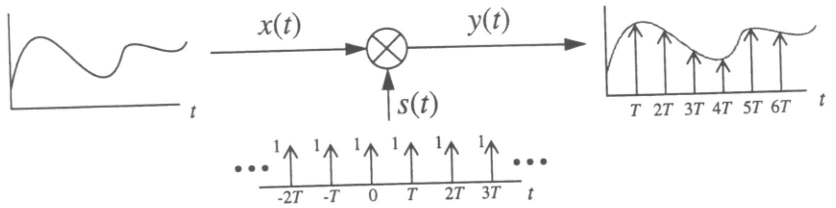
This book:



19.2 Mechanisms of Sampling

Time Domain

Let's examine the sampling process in closer detail, starting in the time domain. The standard paradigm is shown in the following picture. The value of the continuous-time signal is recorded every T seconds by multiplying by an impulse train. The area of the pulses formed in the sampled version is equal to the height of original signal at the sample point.

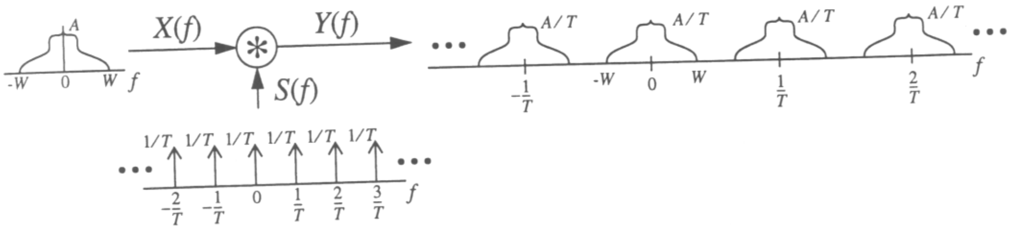


Frequency Domain

Now, let's look at the same process in the frequency domain. First, we assume a bandlimited spectrum for $X(f)$, the Fourier transform of $x(t)$. Remember, a bandlimited spectrum is one where $X(f) = 0$ for $|f| > W$. The reasons for needing a bandlimited spectrum will become clear shortly. The following three facts allows us to determine the spectrum for $Y(f)$, as graphed below.

1. The transform of an impulse train is an impulse train.
2. Multiplication in the time domain becomes convolution in the frequency domain.
3. Convolving with an impulse merely shifts and scales the signal.

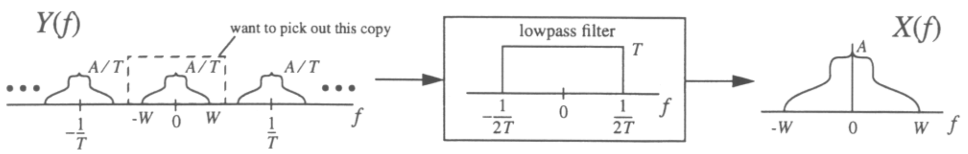
Do not forget to take into account the scale in height of $1/T$ that accompanies the transform of the impulse train.



19.3 The Reconstruction Process

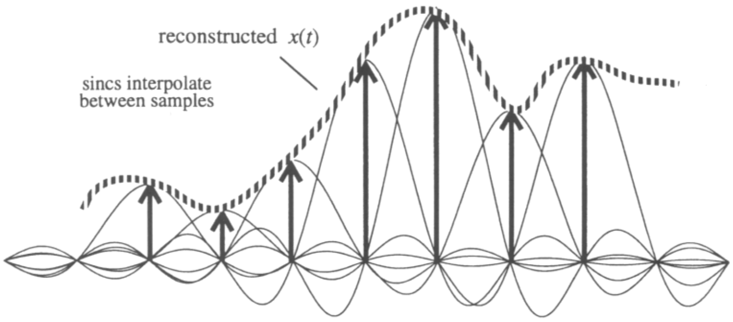
Frequency Domain

The original continuous-time signal $x(t)$ is completely recovered if the original spectrum $X(f)$ can be extracted from $Y(f)$, the spectrum of the sampled signal. By examining the picture of $Y(f)$ shown below, it should be clear that the single spectrum $X(f)$ can be recovered by using an ideal lowpass filter to remove the extra copies. The filter should have cutoffs at $f = \pm 1/(2T)$ and a gain of T.



Time Domain

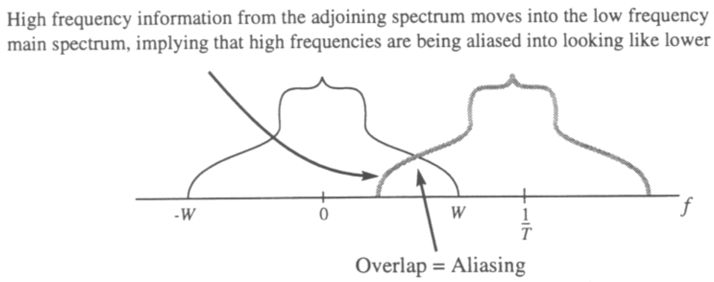
To see how the reconstruction process works in the time domain, recall that a lowpass filter in the frequency domain is a sinc function in the time domain. So, multiplying by a "box" filter in frequency is like convolving with a sinc function in time. Since $y(t)$ is just a series of impulses (samples of $x(t)$), it follows that the reconstructed $x(t)$ is merely the superposition of scaled and shifted sincs!



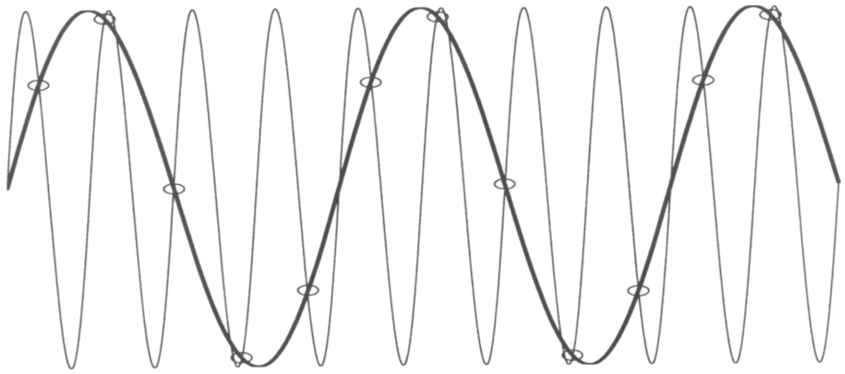
19.4 Aliasing

We have seen that a continuous-time signal can be completely reconstructed from its samples. However, this process will not be possible if the samples are spaced too far apart in time. Why? As the impulses in the impulse train in time are spaced farther and farther apart, the impulses in its transform $S(f)$ get closer and closer together. If the impulses in frequency move too close together, after convolution the copies of the spectra of

$X(f)$ will overlap. Bad! Once the spectrum gets garbled like that, it will be impossible to recover the original continuous-time signal. The situation where spectral overlap occurs is known as *aliasing*.



In the case of overlap, the high frequency information aliases/resembles/looks-like low frequency information. This effect is demonstrated below with a high frequency sine wave that is sampled too slowly. Even though the sine wave is sampled at regular intervals, the reconstruction process fails to reproduce the original signal since the samples are spaced too far apart in time. It is possible to fit a lower frequency sine wave to those same data points, which is a time-domain illustration of how a high frequency signal can get aliased into looking like a lower frequency signal.



19.5 The Nyquist Sampling Theorem

In order to have distortion-free reconstruction of a continuous-time signal, two things have to happen: (1) the continuous-time signal to be sampled must be bandlimited, and (2) the samples must be close enough together in time. From examining the spectrum for $Y(f)$ in Section 19.3, to prevent overlap we must have $1/T - W > W$, or $1/T > 2W$. Since T is known as the sampling period (in seconds), $1/T$ is called the sampling rate (in Hertz). All of this implies the grand result that in order to achieve flawless theoretical reconstruction:

> *The sampling rate must be greater than twice the maximum frequency present in the input signal.*

The critical sampling rate ($2W$) is known as the Nyquist rate. The above statement is known as the Nyquist Sampling Theorem and is definitely one of the most powerful concepts ever seen by an engineering student.

19.6 Practical Considerations

Zero-Order Hold

However nice theory may seem, impulses don't really exist in the real world. In practice, what is commonly done for sampling purposes is known as the zero-order hold. Here, the sampled signal resembles a staircase (i.e. remains flat between sample points). It is still possible to completely reconstruct the original signal; however, the lowpass filter must be slightly modified. See Section 8.1.2 of *Signals and Systems* by Oppenheim et al (1983) for more information. Note that using a standard flat lowpass filter for reconstruction would probably work reasonably well, but it wouldn't be optimal.

Anti-Aliasing Filters

To ensure that aliasing does not occur in the sampling process, people often first pass the signal to be sampled through a lowpass filter in order to insure that it is appropriately bandlimited. This filter is known as an anti-aliasing filter. For example, if you want to sample audio at a relatively low rate of 5KHz, you would first need to pass the sound source through a lowpass filter that had a sharp cutoff at 2.5KHz or lower in order to avoid aliasing of any higher frequency content present in the original signal.

Oversampling

The Nyquist theorem states that it is possible to completely reconstruct a signal as long as you sample at a rate greater than twice the maximum frequency present in the signal. From Section 19.3, if $1/T = 2W$ then the copies of the spectra in $Y(f)$ will be adjacent, i.e. touching. So, in order to reconstruct the original signal you are going to need a very, very sharp analog lowpass filter. Such beasts are not easy to design, so what is commonly done is to *oversample* – that is, to sample at a rate much greater than the Nyquist rate. Then, the copies of the spectra are spread sufficiently far apart so that a more realistic lowpass filter can be used in the reconstruction process. This is why some compact disc players are advertised as having 4x oversampling, 8x oversampling, etc.

A Little Trivia

- The human auditory system can only hear sounds between 20 Hz and 20 KHz.
- Compact discs contain audio sampled at a rate of 44.1 KHz.
- The bandwidth of an ordinary telephone line is about 3KHz (higher frequencies are lost during transmission), which is still large enough to produce an acceptable quality of conversational speech.

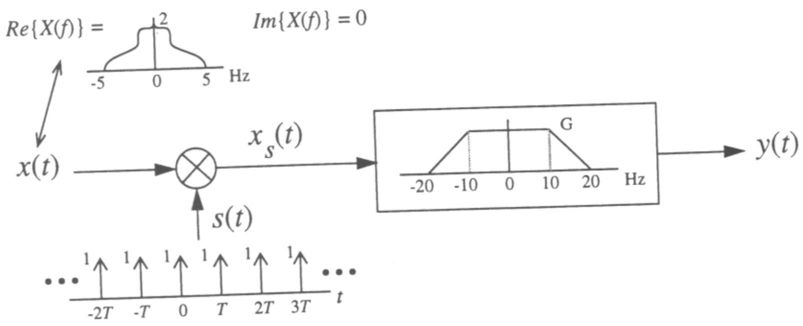
19.7 A Sample Problem (no pun intended)

Questions:

Using the block diagram shown below, answer the following questions:

(a) Sketch $X_s(f)$ for T=20 msec. The real and imaginary parts of the spectrum $X(f)$ are given.

(b) What is the maximum value of T for which the system will produce $y(t) = Kx(t)$ (i.e. perfect reconstruction within a constant factor)?



Answers:

(a) Copies of the spectra of $X(f)$ are replicated at -100, -50, 0, 50, 100, etc. Hertz. Height of each is 100.

(b) Maximum value of T is 40 msec. Cannot sample any slower because of the limitations of the lowpass filter.